

Abdulla Jubair

Kasargod, Kerala, India • +91 8848606337 • abdulla.t.jubair@gmail.com • [LinkedIn](#) • [GitHub](#) • [Portfolio](#)

PROFESSIONAL SUMMARY

AI Engineer with 2 years of hands-on training in Machine Learning, Deep Learning, and NLP/LLM orchestration at Brototype. Experienced in building predictive models (GRU sequence forecasting) and classical classifiers alongside Retrieval-Augmented Generation (RAG) systems and multi-agent workflows (LangGraph). Skilled in Python, data engineering pipelines, and containerized deployments to deliver end-to-end, production-ready AI solutions.

TECHNICAL SKILLS

- **Machine Learning & Deep Learning:** Regression, Classification, Ensemble Methods (Random Forest, XGBoost), Gated Recurrent Units (GRU), Sequence Forecasting, K-Means Clustering, PyTorch, TensorFlow/Keras
- **Generative AI & LLMs:** Transformers, Prompt Engineering (Few-shot, Chain-of-Thought, ReAct), Fine-Tuning (PEFT, LoRA), LangChain, LangGraph (Multi-Agent Systems), Retrieval-Augmented Generation (RAG), LlamaIndex
- **Natural Language Processing (NLP):** GloVe word embeddings, spaCy, NLTK, FinBERT sentiment analysis, Cosine Similarity, CountVectorizer
- **Vector Databases:** FAISS, ChromaDB
- **Programming Languages:** Python (OOP, NumPy, Pandas, Scikit-learn), SQL (PostgreSQL, MySQL, SQL Server)
- **Data Engineering & Cloud Platforms:** ETL/ELT pipelines, Apache Airflow workflows, AWS S3, AWS Glue, AWS Athena, Google Cloud Storage, Google BigQuery
- **BI & Web Analytics:** Power BI Desktop (DAX metrics), Tableau Desktop, Streamlit web applications, Matplotlib, Seaborn
- **DevOps & Software Engineering:** Git/GitHub (Git Flow branching, PRs), GitHub Actions (CI/CD pipelines), Docker containerization, RESTful APIs, FastAPI, Flask, Unit Testing (Pytest fixtures)

KEY AI & DATA SCIENCE PROJECTS

MarketMind AI — Stock Market Intelligence Platform (Capstone Project)

- **GitHub Repository:** [MarketMindAI](#)
- **Summary:** Developed a stock intelligence dashboard offering price predictions, news sentiment tracking, and automated research reports.
- **Bullets:**
- Built a microservices system with 6 FastAPI backend services containerized with Docker and deployed on an AWS EKS (Kubernetes) cluster.

- Created a Gated Recurrent Unit (GRU) model using TensorFlow/Keras to predict 30-day stock prices, and a Random Forest classifier to predict daily price movements.
- Built a daily automated news sentiment tracking pipeline using a PyTorch-based FinBERT model to scrape and analyze stock news.
- Developed a Retrieval-Augmented Generation (RAG) assistant using Sentence-Transformers and FAISS vector databases with custom rules to prevent incorrect answers.
- Created a multi-agent system using LangGraph, linking specialist agents (Risk, Technical, and Sentiment) to write automated market research reports.

Newslyzer — RAG-based News Research Assistant

- **GitHub Repository:** [Newslyzer](#)
- **Summary:** Built an AI news search assistant that extracts text from articles and generates summaries.
- **Bullets:**
 - Created a news extraction app in Streamlit that scrapes and tokenizes text from up to 3 article URLs.
 - Built text chunking and indexing pipelines using MiniLM Sentence-Transformers to create search segments.
 - Designed a local, high-speed vector database store using FAISS to reduce search latency.
 - Set up LangChain RetrievalQA chains to run similarity searches across vector databases.
 - Structured custom templates to query Mistral-7B via the Together API to write summaries with source links.

Fake-News-Detector

- **GitHub Repository:** [Fake-News-Detector](#)
- **Summary:** Developed a machine learning app to classify real and fake news articles using NLP.
- **Bullets:**
 - Developed an NLP classifier using 300-dimensional GloVe word embeddings (en_core_web_lg) to identify fake news.
 - Cleaned text data by removing stopwords and averaging token vectors to create features.
 - Evaluated classical models (Naive Bayes, KNN) and achieved a 99% accuracy rate using K-Nearest Neighbors.
 - Visualized model performance, precision-recall boundaries, and confusion matrices using Seaborn.
 - Deployed the model as a web service using a Flask application to process text submissions and return predictions.

Restaurant-Name-Dish-Planner — LLM Project

- **GitHub Repository:** [Restaurant-Name-Dish-Planner](#)
- **Summary:** Built an AI-driven planner suggesting restaurant names and menu items.
- **Bullets:**
 - Developed an AI planner that suggests restaurant names and customized menus based on user-selected cuisines.
 - Created LangChain PromptTemplates and chains to connect and direct LLM outputs.
 - Integrated Together AI API to query Mistral-7B-Instruct models for text generation.

- Built an interactive Streamlit UI with sidebar filters to display generated lists and formatted menus.
 - Secured API credentials locally by keeping secrets outside version control.
-

EDUCATION

Professional Training – Data Science

Prototype / 2024 – 2026

- Completed hands-on training covering Machine Learning, Deep Learning, NLP, Data Engineering, MLOps, LLMs, Retrieval-Augmented Generation (RAG), Agentic AI, and Cloud deployments.
- Designed and deployed production-ready AI applications using Python, FastAPI, LangChain, Airflow, Azure, FAISS vector databases, Streamlit, Power BI, Git, and RESTful APIs.
- Participated in code reviews and technical discussions, following software engineering best practices and CI/CD pipelines.

Higher Secondary Education (Science Stream)

HSSS Vengad School / 2022 – 2024 / Score: 97%

CERTIFICATIONS

- **Python For Data Science A-Z: EDA With Real Exercises** – Udemy (Nov 2024)
Credential URL: [udemy.com/certificate/UC-a9741356-5d92-43e8-80a1-41dbc7be1459/](https://www.udemy.com/certificate/UC-a9741356-5d92-43e8-80a1-41dbc7be1459/) *Details:* 16 hours of data cleaning, plotting, and preprocessing using Pandas, NumPy, Matplotlib, and Seaborn.
-

LANGUAGES

- **English** (Professional) | **Malayalam** (Native)